PB84-116052

Analysis of Organic Compound Data Gathered during 1980 in Northeast Corridor Cities

(U.S.) Environmental Protection Agency
Research Triangle Park, NC

Aug 83

U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. RECIPIENT'S ACCESSION NO.
EPA-450/4-83-017		PB84 116052

4. TITLE AND SUBTITLE	5. REPORT DATE
Analysis Of Organic Compound Data Gathered During 1980 In Northeast Corridor Cities	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.
Harold G. Richter	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Office Of Air Quality Planning And Standards U. S. Environmental Protection Agency Research Triangle, NC 27711	10. PROGRAM ELEMENT NO. 11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS Same as box 9.	13. TYPE OF REPORT AND PERIOD COVERED 14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

During the summer of 1980, an extensive monitoring project was carried out as part of the Northeast Corridor Regional Modeling Project (NECRMP), the primary purpose of which was to compile an air quality and meteorological data base for use in photochemical models. Since most of these models require ambient concentrations of organic compounds as part of the input information, one of the principal efforts of the study was to collect samples of air from major urban areas of the Corridor and to analyze them for their organic species composition.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS (This Report)	21. NO. OF PAGES
		49
	20. SECURITY CLASS (This page)	22. PRICE

PB84-116052

United States
Environmental Protection
Agency

Office of Air Quality
Planning and Standards
Research Triangle Park NC 27711

EPA-450/4-83-017
August 1983

Air

EPA

Analysis Of Organic Compound Data Gathered During 1980 In Northeast Corridor Cities

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

EPA-450/4-83-017

August 1983

Analysis Of Organic Compound Data Gathered During 1980 In Northeast Corridor Cities

By
Harold G. Richter
Air Management Technology Branch
Monitoring And Data Analysis Division

U.S. ENVIRONMENTAL PROTECTION AGENCY
Office Of Air, Noise And Radiation
Office Of Air Quality Planning And Standards
Research Triangle Park, North Carolina 27711

August 1983

i-a

This report has been reviewed by the Office Of Air Quality Planning And Standards, US EPA, and approved for publication. Mention of trade names or commercial products is not intended to constitute endorsement or recommendation for use.

Analysis of Organic Compound Data Gathered During 1980 in Northeast Corridor Cities

Introduction

During the summer of 1980, an extensive air monitoring program was carried out as part of the Northeast Corridor Regional Modeling Project (NECRMP). The primary purpose of the monitoring program was to compile an air quality and meteorological data base for use in photochemical models. Since most of these models require ambient concentrations of organic compounds as part of the input information, one of the principal efforts of the study was to collect samples of air from major urban areas of the Corridor and to analyze them for their organic species composition.

These data can be used for purposes other than modeling. They can be used to characterize the area surrounding the point of collection, for example (i.e., the emission sources in the vicinity), or they can be used to estimate the contribution of mobile sources to the hydrocarbon burden of the air. To whatever use they might be put, however, the data can contribute to reliable conclusions only insofar as the user knows the qualifications and peculiarities of the data. It is the purpose of this report to acquaint data users with the large body of ambient hydrocarbon compositional data acquired during the NECRMP, to alert the user to limitations in some of the data from certain samples, and to show how the data can be used for certain analyses. Complete data tapes are available as part of the 1980 Northeast Regional Oxidant Study (NEROS) data compilations from J. H. Novak, Meteorology and Assessment Division, Environmental Sciences Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711.

Site Description and Analytical Technique

The Northeast Corridor extends from Washington, DC, in the south to Boston in the north. Samples of air for hydrocarbon analyses were collected at surface sites in each of four metropolitan areas (Washington, Baltimore, New York City-New Jersey, and Boston), as well as above those areas by aircraft. Two surface sampling sites were chosen in each area, one in a high traffic density, center-city location, and the other in an industrial location. Figures 1 thru 4 show the locations of these sites in each metropolitan area.

Two one-hour integrated samples of air were collected in Tedlar and/or Teflon bags at each site on most days of the monitoring program, one from 0600-0700 EDT and the other from 0800-0900 EDT. Aircraft grab samples were collected in stainless steel canisters at various times and locations during the day, as prescribed by a protocol determined by the meteorological conditions prevailing on the individual days. Detailed records were made of the flight paths and the locations, altitudes and collection times of these samples. When time permitted, a third kind of sample was collected. These were called "roving" samples and were grab samples collected at various surface sites downwind of the Washington-Baltimore areas.

Analysis of a sample was carried out as soon as possible after collection, in order to avoid loss of compounds from the containers. The elapsed time was never more than 24 hours. Separate chromatographic and data report systems were maintained for each metropolitan area (although the Washington and Baltimore systems were housed in the same laboratory). Battelle-Columbus Labs collected and analyzed samples in the Boston area, while Washington State University collected and analyzed samples from the other three areas. Analysis

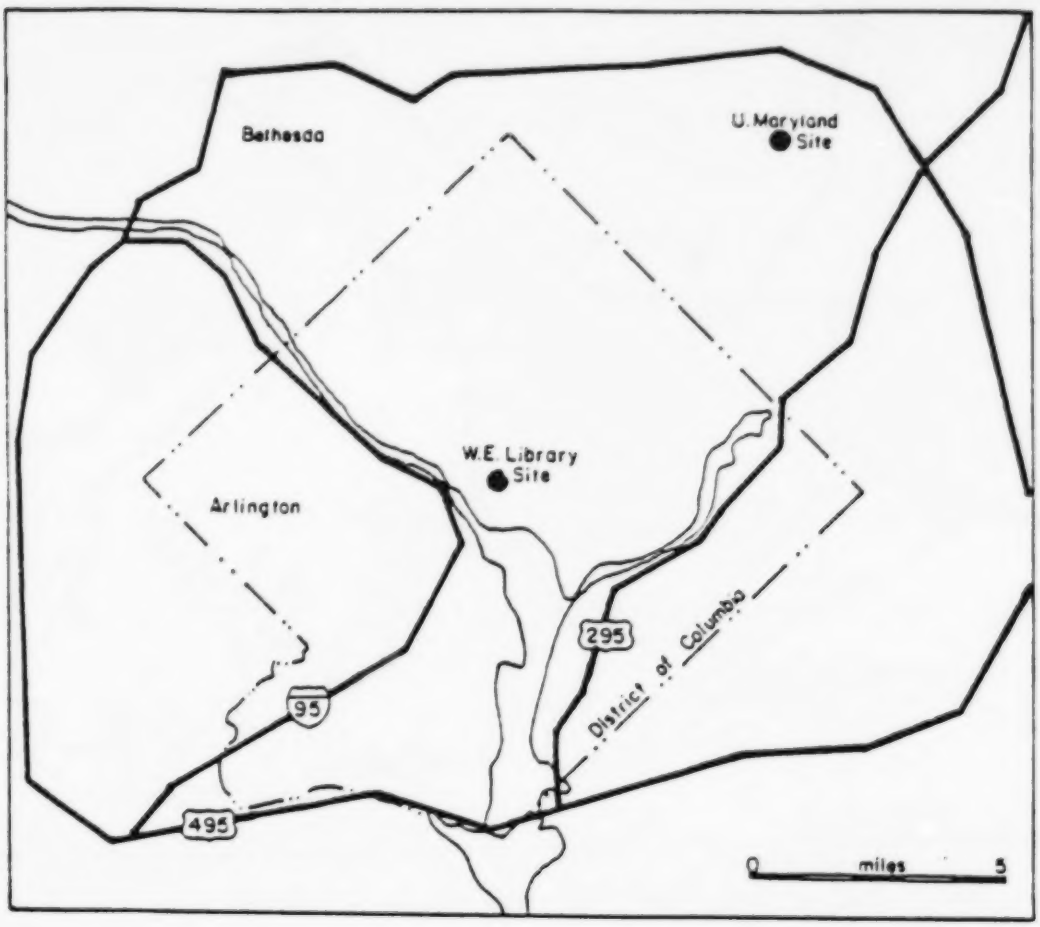

Figure 1 Map showing sampling sites in the Washington, DC area

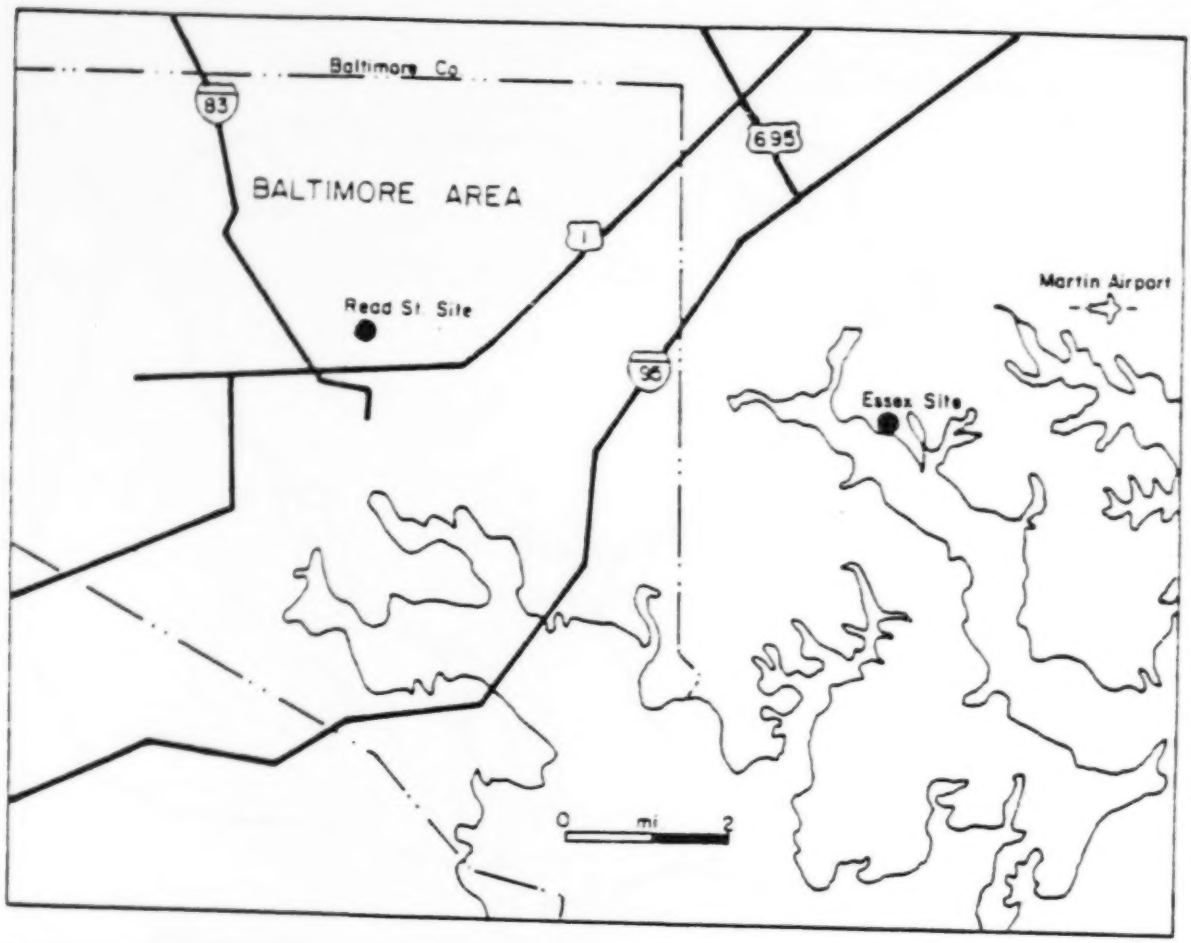

Figure 2. Map showing sampling sites in the Baltimore area.

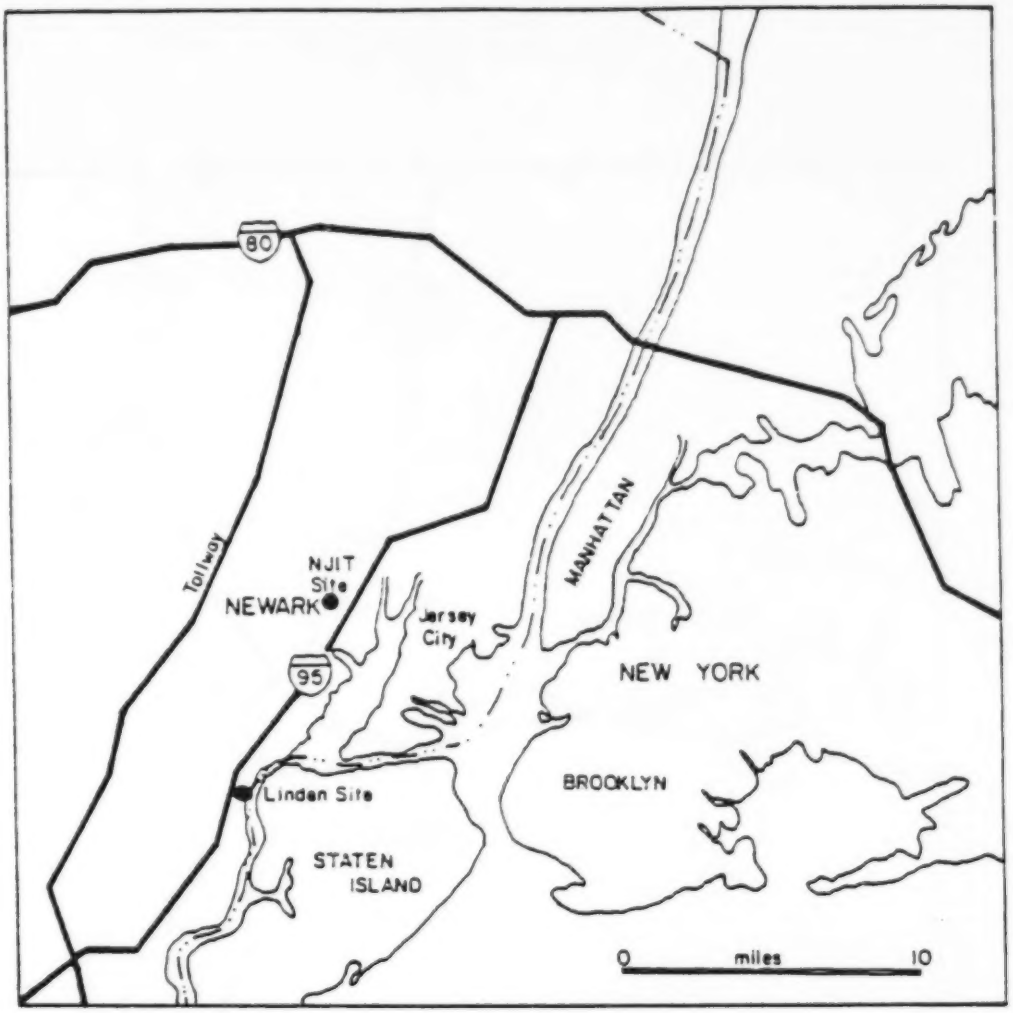

Figure 3. Map showing sampling sites in the New York City area

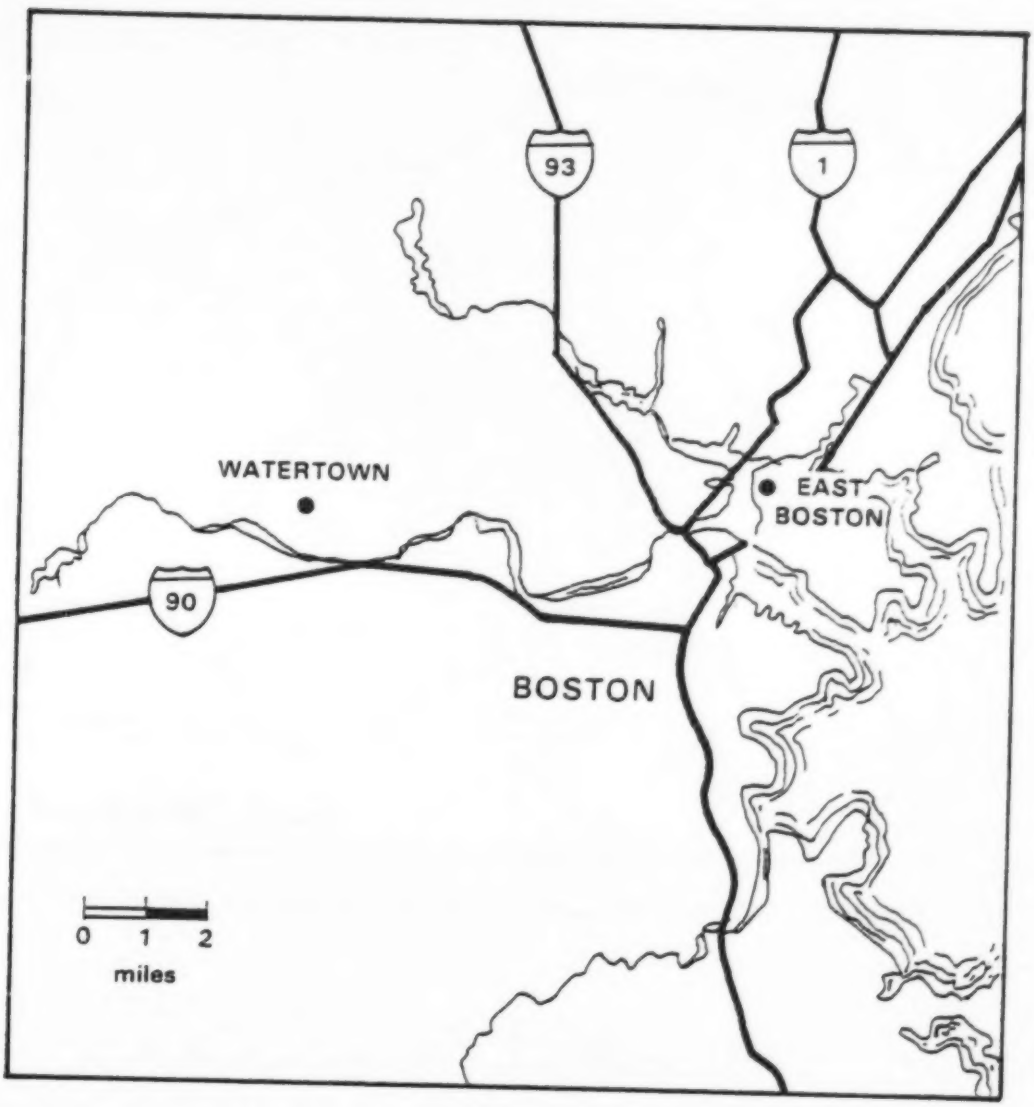

Figure 4. Map showing sampling sites in the Boston area

was by GC-FID (gas chromatography with flame-ionization detection). Details of the columns and procedure have been presented elsewhere.[1-5]

In all, more than 1300 air samples were collected during the 45-day intensive data collection period from July 15 to August 30, 1980. Only 1120 samples were analyzed, however, because of the limited number of samples which could be analyzed in a single day. Table 1 summarizes the number of samples analyzed from each site.

Concurrent with the analysis of ambient samples, an extensive quality assurance/quality control (QA/QC) program was carried out. This included regular analyses of a pentane or neo-hexane (2, 2-dimethyl butane) standard. In addition, each of the four analytical groups had a canister containing a mixture of 14 organic species commonly found in urban atmospheres. For statistical purposes, each group analyzed the mixture at least eight times during the July 15 - August 30 period. Finally, a "round robin" sampling/analysis program was also included in the QA/QC program. This consisted of each analytical group collecting an ambient sample in a stainless steel canister. After analyzing a fraction of the sample, the collecting group then sent the canister to one of the other groups. That group, in turn, analyzed a portion of the sample and sent the remainder to a third group which repeated the operation. In all, five analytical groups participated, and each started at least four samples. [The fifth group was from the Environmental Sciences and Research Laboratory, Research Triangle Park, North Carolina (ESRL/RTP), located in Columbus, Ohio, at the time.]

Table 1

Summary Of Samples From NECRMP Sites

Site	Number of Samples
West End Library, Washington, DC	60
Takoma Park, Washington, DC*	16
University of Maryland	44
Read Street, Baltimore, MD	82
Essex, MD	90
Linden, NJ	81
Newark, NJ	71
East Boston, MA	85
Watertown, MA	90
	619

Aircraft Samples

Location	Number of Samples
Washington - Baltimore	286
New York - New Jersey	66
Boston, MA	86
	438

Roving Samples

Location	Number of Samples
Baltimore, MD - environs	63

* Because of space conflicts, the Takoma Park site in Washington, DC, had to be vacated early in the study period. The University of Maryland site was then chosen to serve as a suburban site for the DC area.

Results and Discussion

A. QA/QC Program

The results of the QA/QC program are reported in detail elsewhere.[6,7] They are summarized here, however, because they are important to the discussion of the data obtained during NECRMP, which follows. In brief, the coefficient of variation for 20-25 analyses of calibration gas (pentane or neohexane) at each laboratory fell between five percent and seven percent. This is very satisfactory. Analyses of the standard mixture of 14 representative organic compounds showed that the individual light compounds (carbon numbers of C-2 to C-5) could be measured with a coefficient of variation of 15 percent or less, but heavier species, particularly aromatic compounds, were subject to much greater errors. Analysis of data from the round robin samples shows that, although the various laboratories generally agreed ($\pm$ 15 percent coefficient of variation) as to the total concentration of compounds present (in parts per billion carbon, ppb C), there was frequent disagreement as to the identity of certain of the minor constituents, as well as sometimes large disagreement (2-fold) in the concentration of total identified olefinic and/or total aromatic compounds present.

These results show that light organic species can be measured with very satisfactory precision. Heavier species are measured with less precision. The reasons for this are not clear at the present time. Some evidence (unpublished) has been gathered in the past, suggesting that passivated stainless steel containers may affect the behavior of heavy species, particularly aromatics, but inconsistent data were obtained from the aircraft samples and the mixture of 14 species during this study to corroborate the point.

From one point of view, the round robin data are very encouraging. They show that skilled analysts can measure ambient hydrocarbon species rather precisely at the ppb C level. Different analysts usually agreed fairly well on the individual concentrations of most of the lighter species (hexane and below) which usually make up about 50 percent of the total concentration. When the data user needs to know only the sum of hydrocarbon species concentrations, he can be confident of a fairly reliable value. Much work needs to be done, however, on quantification of the heavier species, identification of certain unknown peaks, and the frequent irreproducibility of aromatic species data. And, although unrelated to the present discussion, all studies to date of organic compounds in ambient air have almost completely neglected separation and identification of the oxygenated and substituted species known or thought to be present. Whereas these may be in low concentration, a complete characterization of the pollutant burden must include them.

Nevertheless, in spite of the above restrictions, because of the excellent accuracy and precision in measuring the standard sample and the hydrocarbon mixtures, these QA/QC data from NECRMP should allow data users to have confidence in the GC analytical technique and in the data obtained by experienced analysts.

B. Characteristics of the Data

The chromatographic procedure used in analysis of NECRMP samples is capable of separating many compounds in the individual samples. Some of these are identifiable on the basis of retention time. That is, they appear on the chromatogram at times identical to those for known compounds. Frequently,

other peaks appear on the chromatogram at times different from those times at which known compounds appear. These peaks are called "unidentified compounds," and the individual concentrations are summed in order to estimate the total concentration of all organic compounds (excluding methane) in the sample (sum-of-species, ppb C).

During this study, 60 or more compounds were identified in the samples. These are usually divided into three major chemical groups when working with the data: olefins, paraffins and aromatics. One of the first properties of hydrocarbon data sets which many analysts look at is the fractional distribution of total carbon concentration among the chemical groups. These distributions give some information as to the major sources of hydrocarbons influencing the receptor, as well as permitting a rough comparison of data obtained during studies in other cities.

Table 2 summarizes these distributions, as well as listing two other properties of the samples which warrant discussion: anomalous values of C_2H_4 (ethylene) and the fraction of unknowns in the samples. First, however, the reader may notice that the number of samples from each site listed in Table 2 may be slightly lower than the number listed in Table 1. This is due to the fact that duplicate samples were collected occasionally, and their analyses are not included in the calculations of Table 2.

The percentages of the species were calculated by summing the concentrations of all members of the group (e.g., olefins) and dividing by the sum of all identified species concentrations. (The difference between the sum of these three percentages and 100 percent at any given site is the percent of acetylene in the sample.) From the standard deviations, it is seen that the

Table 2

Characteristics Of Data From Surface Samples

Site	Number of Samples	Average Percent Olefins*	σ	Average Percent Paraffins	σ	Average Percent Aromatics	σ	Average Percent Unidentified	C_2H_4 Anomalies # of Samples
West End Library, Washington, DC	58	9.1	6.1	58.7	5.6	28.8	7.0	17.5	36
Takoma Park, Washington, DC*	16	5.8	5.6	63.0	11.6	28.0	9.5	19.4	12
University of Maryland	43	9.2	2.9	54.0	8.8	33.8	9.7	18.4	2
Read Street, Baltimore	80	10.6	5.3	59.1	7.3	27.5	7.7	12.3	27
Essex, MD	89	6.2	2.8	60.6	7.9	30.9	8.4	15.2	51
Linden, NJ	71	8.7	3.6	64.3	11.0	25.4	12.3	12.1	42
Newark, NJ	71	9.3	3.0	57.5	5.7	30.3	6.9	14.4	15
East Boston, MA	85	8.4	3.7	65.7	10.5	23.2	7.4	15.4	--
Watertown, MA	90	14.2	6.2	58.3	7.7	24.3	5.7	18.7	--
Average - all sites	603	9.0	4.5	60.1	8.5	28.0	8.3	15.9	--

* Calculated from data for all samples, including those samples for which ethylene concentrations were recorded as zero. See discussion on C_2H_4 anomalies.

-12-

samples vary appreciably in their fractional composition of some of the groups, particularly the olefins. The paraffin fraction, on the other hand, is much less variable (coefficient of variation) among samples from a given site than are the olefins or aromatics.

The penultimate column of Table 2, called "Average Percentage of Unidentified," presents a characteristic of the data which should be of interest to data users. As described above, unidentified compounds are those which appear on the chromatogram at retention times different from those times at which known compounds appear. Little or nothing is known about the chemical structure of these compounds; thus, they cannot be placed in the olefin, paraffin or aromatic categories. If the data user is interested in knowing the total concentration of olefins (or paraffins or aromatics) in a sample, he must consider the probability that some of the compounds in the unidentified fraction are olefinic. That is, the total concentration of olefins calculated from a simple summation of the identified olefins in a sample is less than actually existed in the sample as collected. Without more information about the distribution of hydrocarbon species in ambient air, it would seem that the only alternative a data user has is to distribute the unidentified compounds among the three major groups in the same proportion as these individual identified groups stand to the total identified hydrocarbon concentrations in the sample. Put more simply, if the concentration of olefins were found to be "X" percent of the concentration of identified compounds in a given sample, then the data user should assume, as a first approximation, that the fractional concentration of olefins in the unidentified compounds is also "X" percent. Whereas this will not change the

fractional relationships among the major groups, it will change the absolute concentrations of the group sums, as well as the calculated group fraction of the recorded sum-of-species. Table 3 shows how these concentrations change, using data from an actual sample (Read Street, Baltimore, 8/13/80, 0600-0700 EDT).

Comparing columns 2 and 6, or columns 3 and 7, shows that there is a 25 percent increase in the recorded concentration (or fraction) of the group sums if the unidentified compounds are distributed proportionately among the three groups. Acetylene, of course, remains the same concentration and fraction, since it is an identified species. Such an increase in group concentrations would very probably have an appreciable impact on some uses of the data.

The last column of Table 2, "C_2H_4 Anomalies," requires discussion, too, since it may be of importance for some data uses. After analysis of the samples, the analysts reported a zero value for ethylene concentration in each sample for which there was evidence of ethylene contamination. Whenever this contamination is found, it is almost always in samples from monitoring stations housing chemiluminescent ozone analyzers. The reported sum of olefins in samples which were contaminated is, therefore, much too small, if only the concentrations of identified olefins are summed, since previous studies[3,9] have shown that ethylene frequently makes up about half of the total olefin concentration in an uncontaminated sample. Table 2 shows that many samples seem to have been contaminated with ethylene in some of the stations.

The problem with ethylene contamination affects other aspects of the analysis of hydrocarbon data, although the impact may be less than it is when examining only the olefin data. The total sum-of-species concentration,

Table 3

Summary Of Data From Surface Samples

Read St., Baltimore 0600-0700 EDT, 8/13/80

Group	ppb C in Sample	Fraction of Total	Fraction of Identified (less acetylene)	ppb C Unidentified (fraction x 123.8)	Corrected ppb C	Corrected Fraction
Olefin	53.5	0.083	0.106	13.1	66.6	0.104
Paraffin	306.5	0.478	0.606	75.0	381.5	0.594
Aromatic	145.5	0.227	0.288	35.7	181.2	0.283
Acetylene	12.5	0.019	-	-	12.5	0.019
Identified	518.5	0.807	-	-	-	-
Unidentified	123.8	0.193	-	-	-	-
Total	641.8	1.000	-	-	641.8	1.000

as recorded, is less than it actually was in the sample by the amount of the missing ethylene concentration. This will typically cause an underestimation of the sum-of-species concentration by about five percent, which may be relatively unimportant for most uses of the data, but it may be important for some. It will affect the calculated paraffin and aromatic fractions (of total hydrocarbon concentation) to some extent also.

If the data user needs to estimate the value of a missing ethylene concentration, he has two alternatives. He can either assume the missing concentration is equal to that of the sum of all other olefin concentrations, or he can calculate the average C_2H_4/C_2H_2 (ethylene/acetylene) ratio in uncontaminated samples from the same sampling site and use that ratio and the acetylene concentration in the sample of interest to approximate the missing ethylene concentration. Both methods usually give similar concentrations, although the latter method is probably more correct.

A different anomaly in ethylene concentrations is noted in data from samples collected and analyzed at the two Boston sites. In many of these samples, ethylene is found to make up 70-90 percent of the total olefin compounds. This cannot be attributed to a source of ethylene near the Boston sites, however, as shown by an inspection of data from the round-robin samples. (Round-robin samples were grab samples -- not 1-hour integrations -- collected usually in the vicinity of the monitoring stations and were thus similar in species concentration to the daily samples.) In those round-robin samples collected in Boston, and analyzed by the other laboratories, ethylene made up about 50 percent or less of the total olefin concentration, similar to that found in samples from other areas. Furthermore, a close analysis of

the round-robin data shows that the Boston analysts reported ethylene concentrations very similar to those concentrations found by other analysts of the same samples. It thus seems that the Boston analysts reported fewer or smaller concentrations of other olefins in the samples, thereby making the ethylene/total olefin ratios high. Whatever the reason, these developments introduce uncertainty as to the reliability of the olefin group fractions and, to a lesser extent, of other group fractions and sum-of-species concentrations for the Boston sites.

C. Uses of the Data

In addition to using the NECRMP hydrocarbon species data for modeling purposes, there are other uses which make the data base valuable, in spite of its limitations. Some of these are discussed below. Other uses and analyses will undoubtedly be made in the future.

First, it is necessary to state, explicitly, a rule which all data users should follow when using these data, as well as when using all other data bases: Simple summaries of data may be very misleading. Data users should become very familiar with the details of the data base they work with. Whereas it may be time consuming, it is imperative that users examine individual sample data, rather than summaries, in order to avoid arriving at wrong conclusions. The rule seems obvious, but it is too frequently ignored.

The value of this rule should be clear from the discussion above about anomalies in olefin concentrations in this data base. Table 4 shows how a simple summary of average olefin fraction (all samples) at the various NECRMP sites differs from a summary of average olefin fraction of only those samples in which ethylene concentrations were reported. Large errors in calculated

Table 4

Comparison of Olefin Fractions

-- All Samples Vs. Only Those Reporting Ethylene --

Site	All Samples			Only Those With C_2H_4		
	mean	σ	n	mean	σ	n
West End Library, Washington, DC	0.091	0.061	58	0.139	0.048	22
Read Street, Baltimore, MD	0.106	0.053	80	0.122	0.053	53
Essex, MD	0.062	0.028	89	0.077	0.033	38
Linden, NJ	0.087	0.036	71	0.100	0.035	29
Newark, NJ	0.093	0.030	71	0.097	0.026	56
			369			198
East Boston, MA	0.084	0.037	85	0.084	0.037	85
Watertown, MA	0.142	0.062	90	0.142	0.062	90

olefin fractions are thus possible at some of the sites if the ethylene problem is unknown or ignored. Olefin fractions at the two Boston sites are included in Table 4 for comparison with the other sites. It is apparent that a significant difference in the spectrum of hydrocarbon species exists among the several sites, although more data are necessary in order to determine how statistically important this is.

1. Compound/Acetylene Ratios

One use to which hydrocarbon species data is sometimes put is to estimate the contribution of automotive emissions in the samples. This is done by assuming - and it is usually a sound assumption - that the only source of acetylene is from automobiles. If one knows the fraction of acetylene in auto exhaust, then the olefin/acetylene ratio or the paraffin/ acetylene ratio in an ambient sample is a measure of automobile contribution to the hydrocarbons in the sample. These fractions have been measured in both pure exhaust[8] and in highway tunnels.[3,9,11] Table 5 summarizes these NMHC/acetylene ratios.

There is some agreement among ratios, but the contribution of auto exhaust to observed hydrocarbon levels in ambient samples cannot be asserted too rigidly. Using Lonneman's ratio from the recent Lincoln Tunnel study for total NMHC/acetylene of 29.7, which is believed to be the most representative of exhaust from the current mobile source fleet, the percent of automobile contribution to observed hydrocarbon concentrations at the various NECRMP sites can be estimated by multiplying the average acetylene concentration by 29.7 (giving an estimate of the HC sum-of-species concentration contributed by mobile sources) and dividing that product by the average HC sum-of-species concentration at the site. Table 6 shows these

Table 5

Nonmethane Hydrocarbon/Acetylene Ratios in Auto Exhuast

Ratio	Black[8] (exhaust)	Lonneman[3] (tunnels)	Lonneman[9] (Houston)	Lonneman[11] (Lincoln Tunnel)
Olefin/Acetylene	4.7	3.24	3.08	6.05
Paraffin/Acetylene	10.2	6.81	7.59	13.1
Aromatic/Acetylene	3.0	3.87	5.22	9.53
Total NMHC/Acetylene	22.0	13.90	15.90	29.7

Table 6

Estimated Automobile Contribution To Observed HC Sum-of-Species

Site	Acetylene ppb C, average	HC Sum-of-Species (ppb C, average)	Percent Auto
West End Library, Washington, DC	17.8	714	74
University of Maryland	16.5	666	74
Read Street, Baltimore	15.6	639	72
Essex, MD	11.1	606	54
Linden, NJ	10.6	961	33
Newark, NJ	12.6	510	73
East Boston, MA	12.1	777	46
Watertown, MA	7.7	352	60

results.

At four of the sites, one may conclude that more than 70 percent of the observed hydrocarbons come from auto exhaust. The Linden, New Jersey, site shows a smaller contribution from automobiles. However, there is a large refinery nearby, and a more detailed analysis of the data would be required to establish the impact from this source more firmly. Similarly, a petroleum tank farm was located upwind (on certain days) from the East Boston site. On these days, the total hydrocarbon concentration in the samples was several ppm C, most of it paraffinic. Excluding these data from the calculation would increase the estimated mobile source contribution to nearer 70 percent. This higher number for automobile contributions to ambient hydrocarbons would then reflect the situation which prevails around the sampling site most of the time.

There is no rigorous proof of the correctness of this method for estimating the fraction of mobile source contribution to ambient hydrocarbons. The basis seems sound, and the calculated fraction agrees with one's impression of mobile sources in the neighborhood, but an experimental program to test the the assumptions would be very difficult to mount.

Another example of compound/acetylene ratio analysis can show how this technique may help characterize the vicinity around a sampling site, as well as to illustrate how the use of data summaries without a critical review may be misleading.

Some hydrocarbon compounds in the ambient air usually come from a limited number of sources. Large concentrations of ethane, for example, probably have commercial natural gas as their source. Large concentrations of propane probably come from liquefied petroleum gas sources. Benzene and toluene generally come from evaporated solvents, while n-butane and i-pentane

come from gasoline vapors. Since the absolute concentrations of any of these species may vary from site to site (because they also are found in auto exhaust), the best way to compare one site with another is by comparing the compound/acetylene ratios (i.e., acetylene is a normalizing factor).

The average n-butane/acetylene and i-pentane/acetylene ratios in atmospheres not dominated by point sources of these two compounds is about 3 and 5, respectively.[13] These ratios are found at all of the NECRMP sites, except East Boston and Linden, where they were 12.5 and 15.7, respectively. This indicated a probable source of gasoline evaporative loss in the vicinity of the site. Closer examination of the individual sample data for the East Boston site, however, showed that in only a few samples were these ratios elevated. All others were in the customary range. On those certain days, the ratios were very high (100-300). These dominated the averaging process, potentially leading to an erroneous conclusion that the site was continually surrounded by such a source.

Examination of wind directions on those days of high ratios showed that the wind was always from the NW, 310°-350°. The sum-of-species concentrations on these days always exceeded 1 ppm C. The conclusion seems now very clear: there was a large source of butanes and pentanes, probably from evaporative emissions, NW of the East Boston site. On further investigation, it was found that a gasoline storage tank farm was located at some distance NW of the site. In retrospect, the choice of the location for the monitor proved to be a less than ideal location at which to measure hydrocarbons. However, this was suspected when the siting decisions were being made, but other factors dictated locating the hydrocarbon monitor at this site.

2. Sum-of-Species/NMOC Comparisons

Another use to which species data are put is to compare sum-of-species concentrations with concentrations of total NMOC recorded by collocated continuous instruments. At seven of the sampling sites (Read St., Baltimore; Essex, MD; West End Library, DC; Linden, NJ; Newark, NJ; East Boston; and Watertown, MA), continuous NMOC analyzers (MSA 11-2) were set up to measure ambient hydrocarbons during the NECRMP. These insruments were calibrated with great care (propane standard) and maintained by experienced technicians in order to gather the best data possible. They sampled ambient air from the same manifold which supplied air to the Teflon/Tedlar bags for species analysis. Analysis of data from these simultaneous hourly measurements may be interpreted in more than one way: a comparison of how well (or poorly) behaved are continuous NMOC analyzers compared with the sum-of-species; how comparable are the data from different instruments of the same make; or how confident a modeler might be in using NMOC instrument data for ozone modeling.

Figures 5 thru 11 show the results of comparisons of NMOC data from analyzers with sum-of-species. Data have been plotted on log-log scales in order to simplify comparison of the different sites. Such plots provide greater resolution at the lower end of the scale than do linear plots, as well as allowing percentage differences to be read directly from the plot. Sum-of species includes the unidentified compounds, as well as the identified species. As noted previously, no correction was made to the sum-of-species data for missing ethylene concentrations in those samples where ethylene contamination was evident (and the ethylene concentration was recorded as zero). To do so would increase all sum-of-species concentrations by a few percent, thereby translating some of the points in the figures slightly to the right.

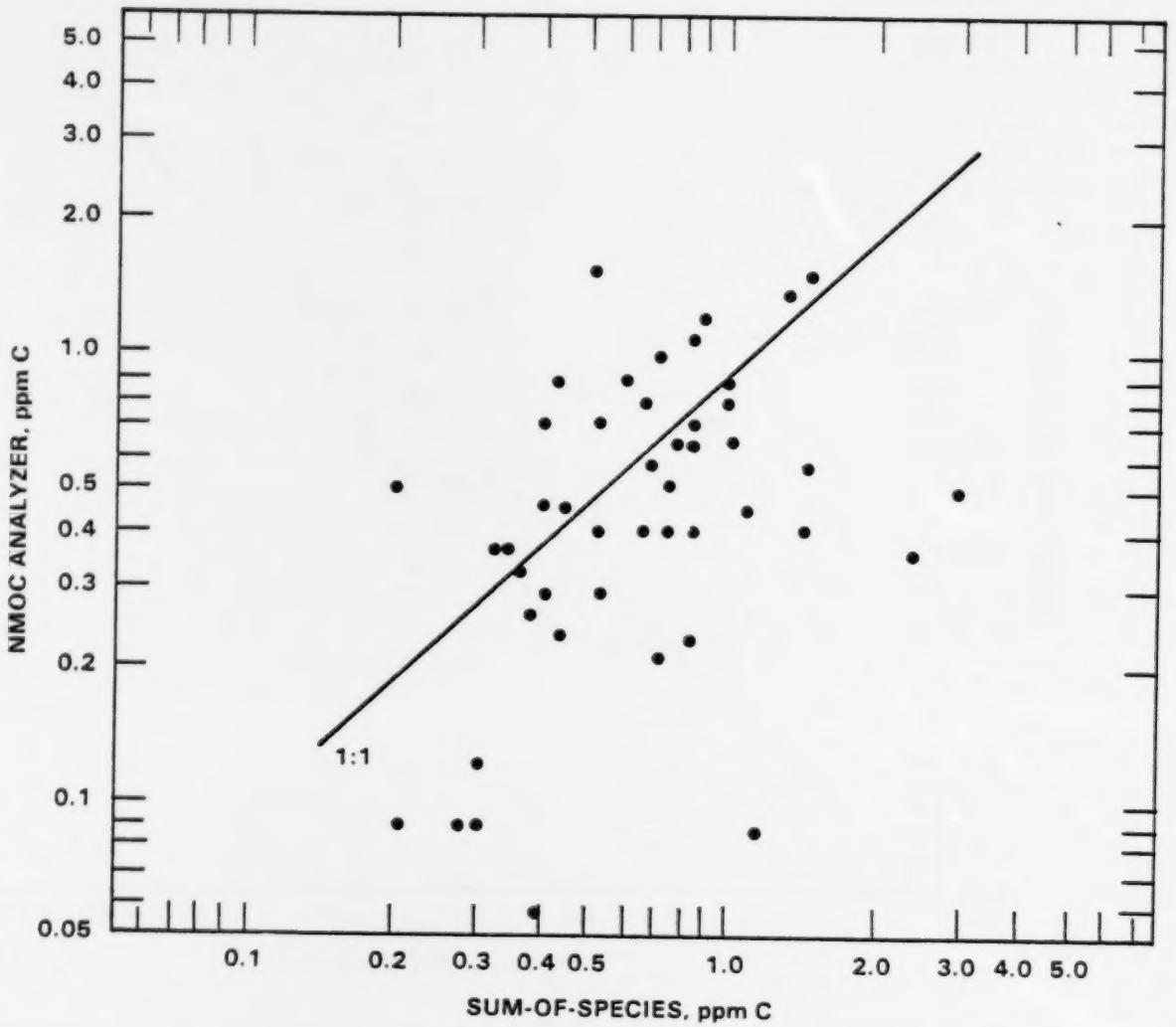

Figure 5. West End Library: NMOC Analyzer - Sum-of-Species Comparison

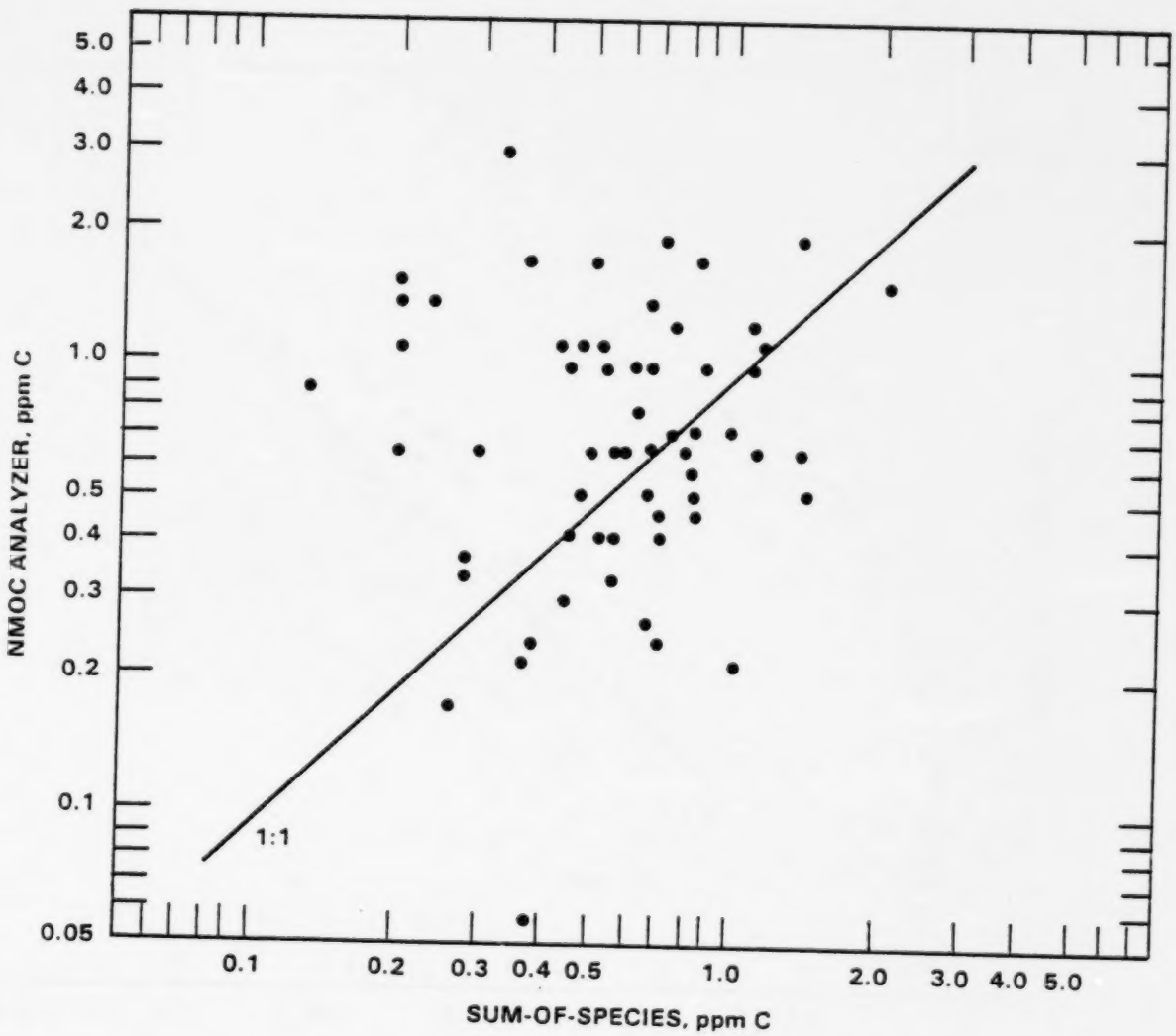

Figure 6. Read Street, Baltimore: NMOC Analyzer - Sum-of-Species Comparison

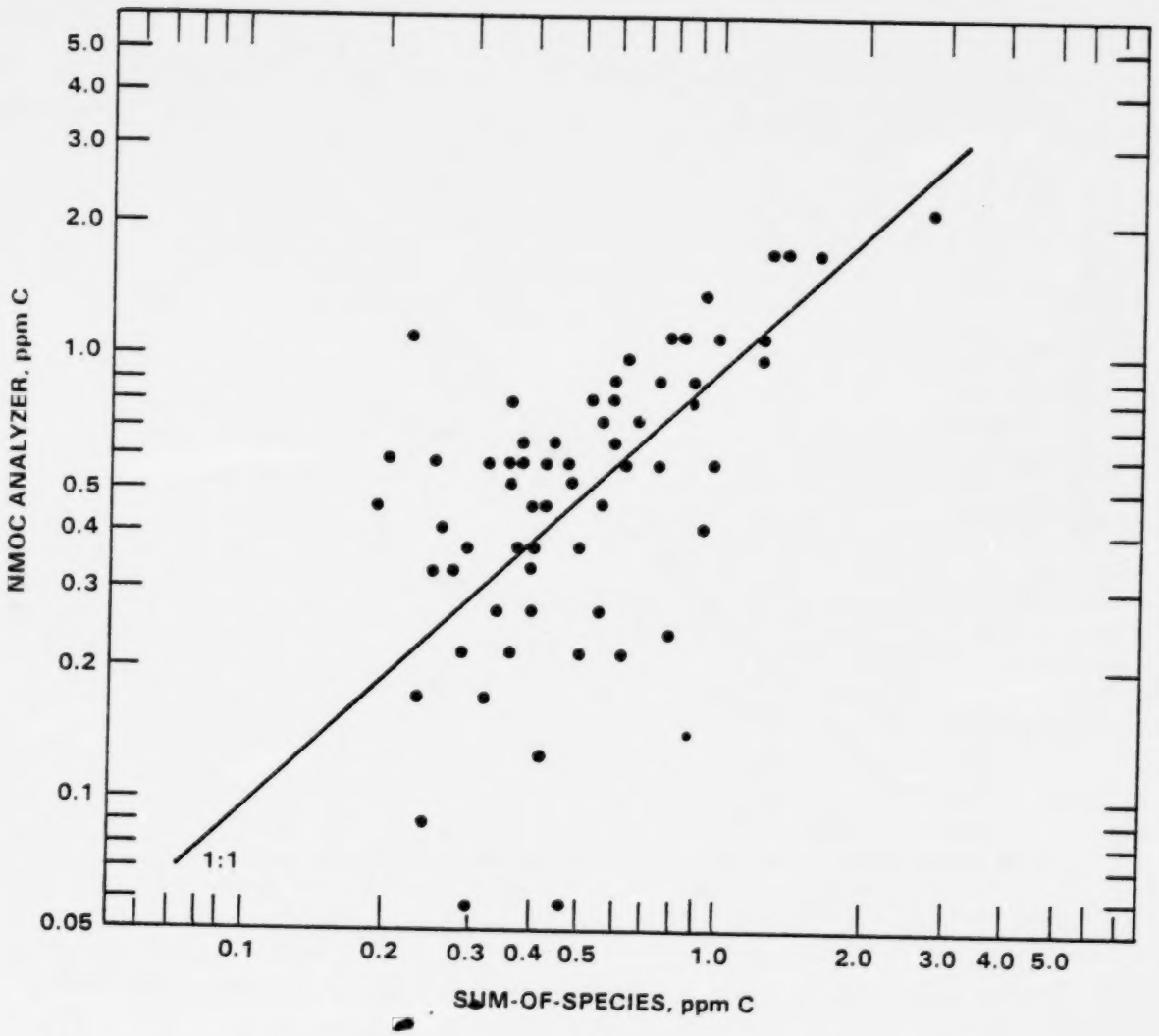

Figure 7. Essex, MD: NMOC Analyzer - Sum-of-Species Comparison

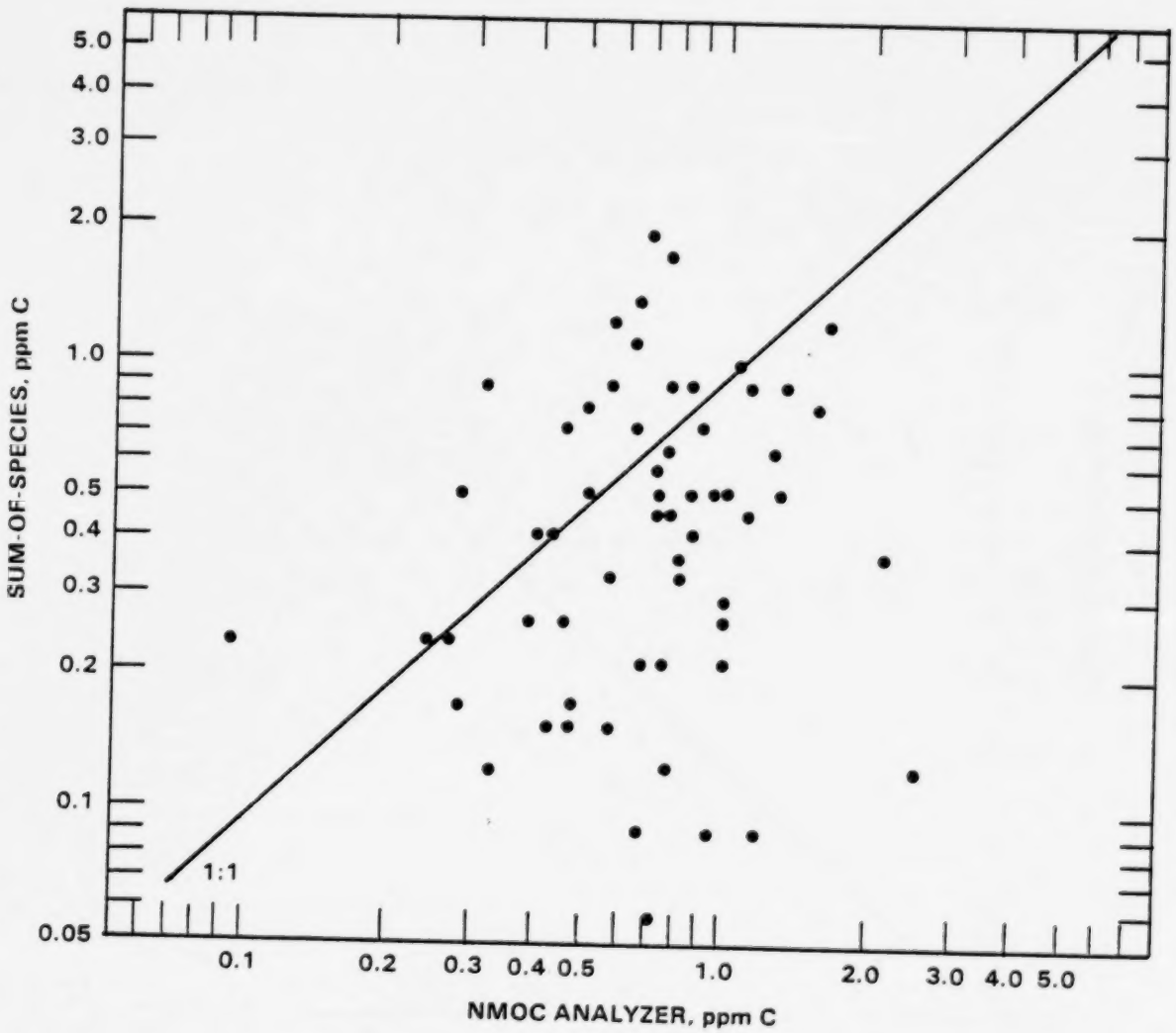

Figure 8. Linden, NJ: NMOC Analyzer - Sum-of-Species Comparison

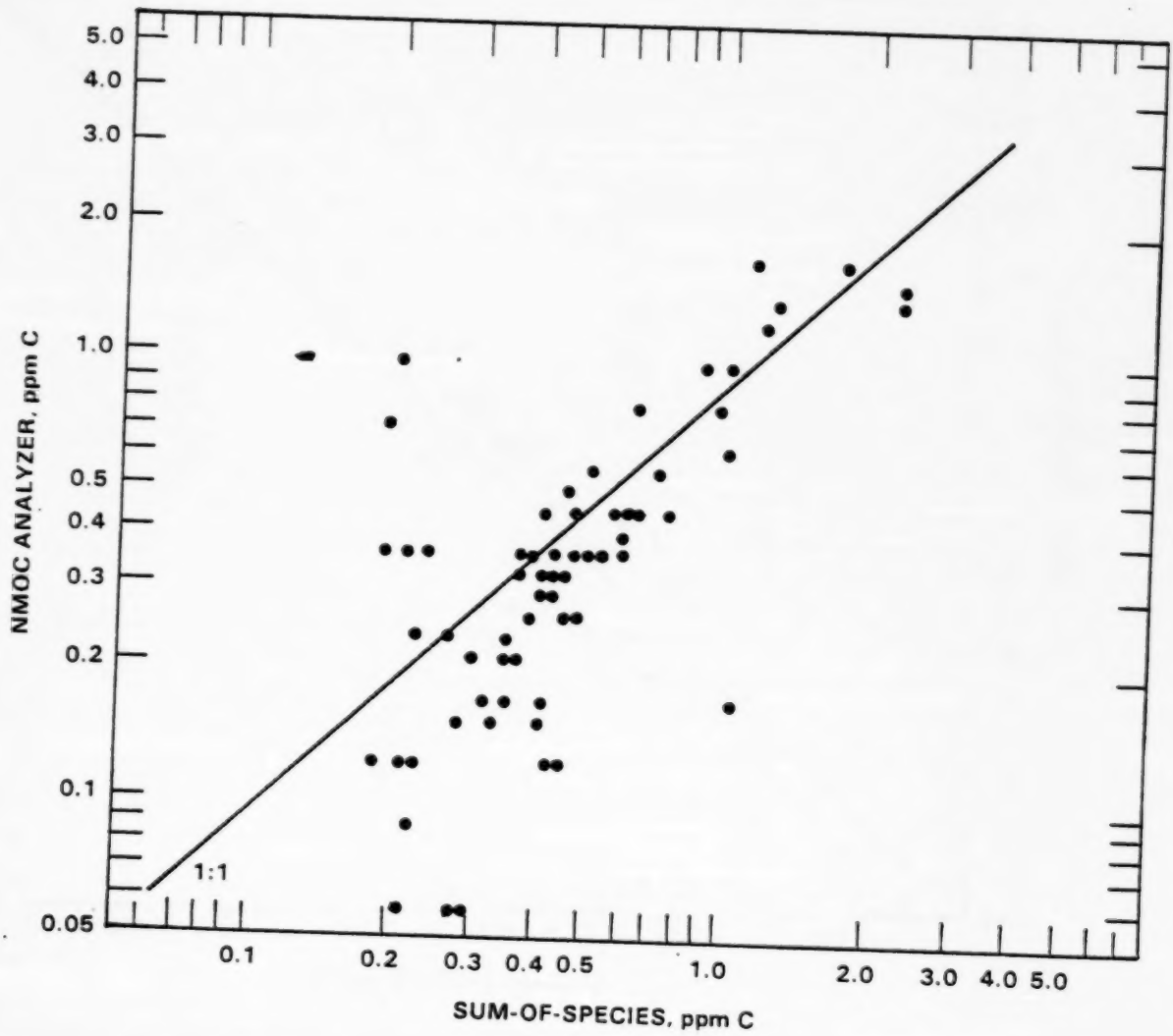

Figure 9. Newark, NJ: NMOC Analyzer - Sum-of-Species Comparison

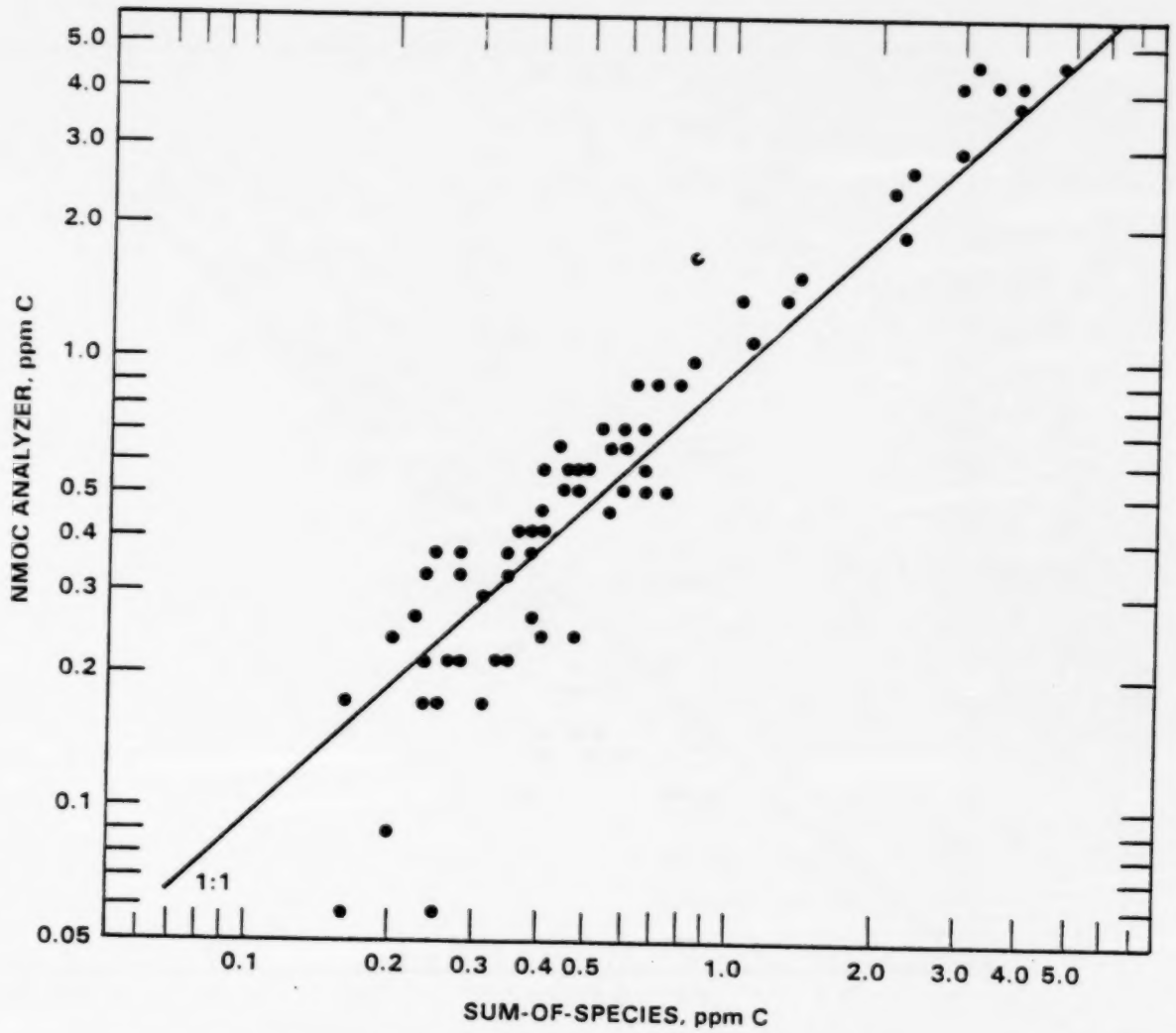

Figure 10. East Boston, MA: NMOC Analyzer - Sum-of-Species Comparison

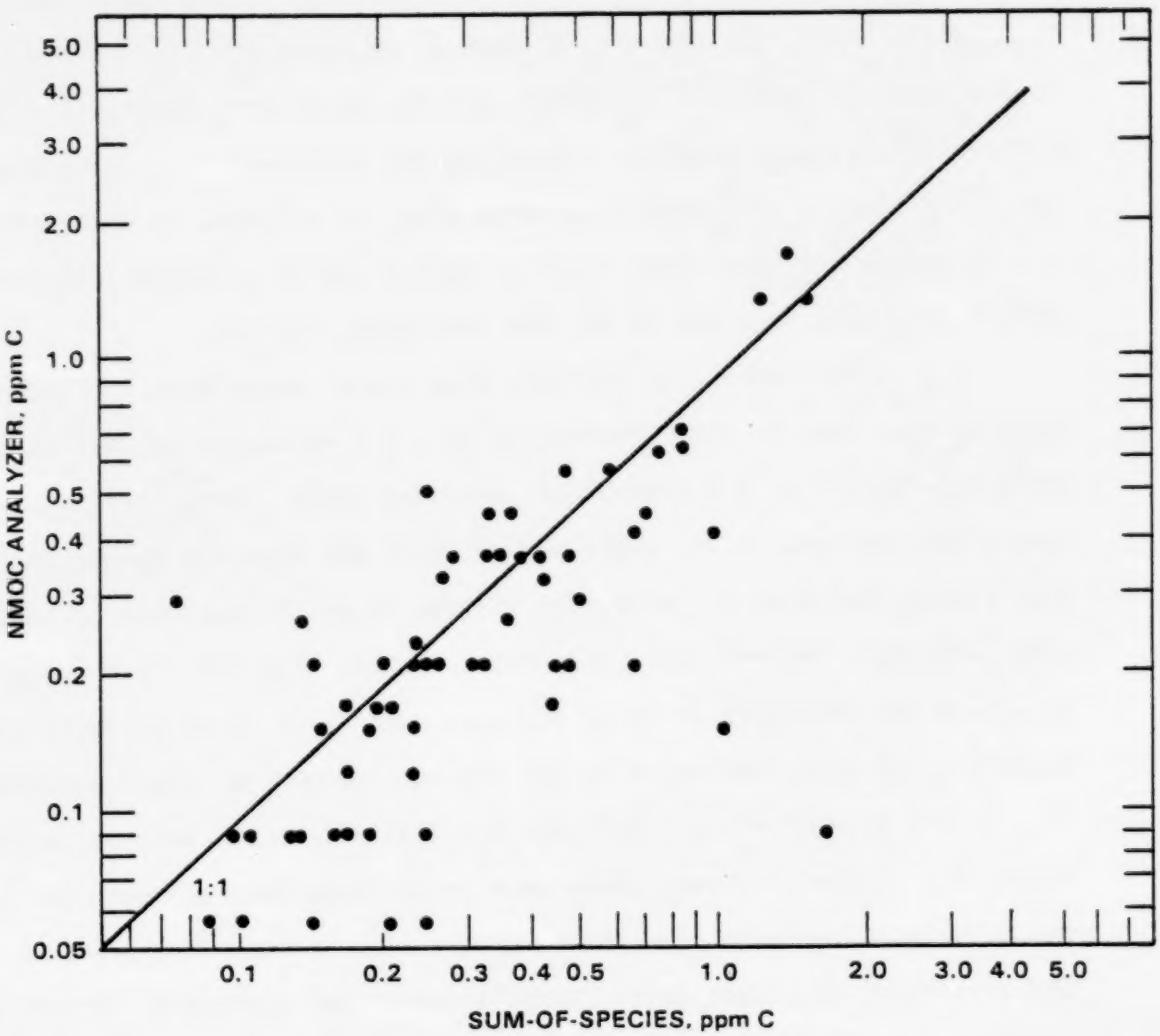

Figure 11. Watertown, MA: NMOC Analyzer - Sum-of-Species Comparison

Perfect agreement between all of the measurements would have all of the data lying on the 1:1 line. There is no absolute way to determine which of the measurements in such a comparison is correct. There is some justification, too, for thinking that the two procedures do not measure equally well the same species in all the samples. Because of the known difficulties with continuous NMOC instruments,[10,12] however, and the satisfactory results which the five analytical groups obtained in analyzing the round-robin samples (± 15 percent coefficient of variation in sum-of-species, 65 analyses), it seems reasonable to assume, for this study, that the sum-of-species procedure yields more uniformly reliable data than do the NMOC instrument readings.

Little can be generalized about these comparisons. It must be borne in mind that the NMOC instruments were not maintained as well as they would have been in a well controlled laboratory study. Thus, there could be some slight inaccuracies in sample timings (both NMOC data and species collection timing), calibration, voltage and/or temperature fluctuations, or perhaps other problems. Nevertheless, the data obtained from the instruments is thought to be the result of above average attention to technical details and to be an accurate representation of the data obtainable from these instruments.

At most of the sites, the data scatter widely. Only at the East Boston site (Figure 10) does there seem to be consistent agreement between the two methods, although there is a positive bias of about 15 percent in the NMOC continuous instrument data. Table 7 summarizes statistical information about the least squares analysis of these natural logarithmic data. As the table shows, only data from the East Boston site exhibit a high r-square value.

At five of the sites, half or more of the samples contained less than 600 ppb C of hydrocarbons. This is of some concern, since the comparisons

Table 7

Sum-of-Species/NMOC Comparisons, Least Squares Regressions

Site	Slope	Intercept	Standard Error	r-square
West End Library, Washington, DC	0.552	- 0.552	0.672	0.169
Read Street, Baltimore, MD	0.113	- 0.283	0.713	0.0077
Essex, MD	0.835	- 0.101	0.599	0.354
Linden, NJ	0.531	--	0.865	0.141
Newark, NJ	0.987	- 0.277	0.574	0.467
East Boston, MA	1.108	+ 0.095	0.327	0.887
Watertown, MA	0.750	- 0.568	0.574	0.475

of NMOC data and sum-of-species show that there can be large differences in concentrations, as measured by the two techniques. Data from the NMOC analyzers were all carefully examined by the contractor responsible for their operation. All questionable data were discarded; thus, only "verified" NMOC data - that which might be used by an agency for its regulatory purposes have been archived. There is good reason to expect no more accurate data from NMOC instruments in routine field monitoring programs which are likely to be maintained with less care than were the instruments in this study. The potential for large error in hydrocarbon concentration data is then reflected in greater uncertainty in modeling predictions, if the continuous instrument data were to be used for this purpose.

The NMOC instruments should not have been affected by ethylene contamination, as were the samples destined for species analysis, for the following reason. Most, if not all, of the collocated ozone analyzers were fitted with ethylene oxidizers which were designed to destroy the excess ethylene after it left the ozone instrument. Thus, assuming properly functioning oxidizers, no ethylene should have been exhausted to outside where it could be entrained in manifold air from which all instruments (including both the Tedlar bags and the NMOC device) sampled. Moreover, if ethylene had entered the NMOC instruments through entrainment, concentrations of NMOC from these monitors should have been surprisingly large, most if not all the time, since the ozone and NMOC instruments were running continuously. On the other hand, leakage of ethylene at the instrument and from plumbing connections inside the monitoring shelters could have exposed the Tedlar bags to about 24 hours of a low level (ppm) ethylene atmosphere. Small molecules, ethylene particularly, are known to diffuse readily through thin Tedlar membranes. No

such diffusion is possible into the NMOC instrument. This is not a completely satisfying explanation, because it is difficult to believe, with the air conditioning maintained in the sampling stations, that much ethylene could collect and diffuse into the bags. In summary, the source of the ethylene contamination of GC samples at certain times is unknown.

In a laboratory study[10] comparing continuous NMOC analyzer data with sum-of-species data, smaller scatters of data were observed. In that referenced study, more than 80 percent of the NMOC data from the MSA 11-2 instrument lay above the 1:1 line. The slope was 1.06 and intercept 0.079, $n = 48$, and $r^2 = 0.951$. These studies show that much more work is needed to resolve the differences or to get better, consistent agreement between these two measurement techniques.

3. Sum-of-Species/NO_X Ratios

At some of the sites, NO_X chemiluminescent analyzers were used to measure ambient NO_X concentrations. Not as many data were collected from these instruments as was desired because of malfunctions. Table 8 shows the average sum-of-species/NO_X ratios. The hourly sum-of-species averages have been listed for all sites studied, as well as the (simultaneous) hourly NO_X averages. (Note that the sum-of-species/NO_X ratio recorded for the sites in Table 8 is not obtained by dividing the average sum-of-species by average NO_X. The ratios in column 3 were obtained by averaging all the individual sum-of-species/NO_X ratios recorded at those sites.)

Except as described below, all data pairs were used for these averages, including those hydrocarbon samples which exhibited C_2H_4 anomalies. Correction of these samples for missing C_2H_4 concentrations would increase the ratios of Table 8 by less than five percent.

Table 8

Sum-of-Species/NO_x Ratios

Site	Sum-of-Species (Average)	NO_x (Average)	Ratio (Average) ± Standard Deviation	n
West End Library, Washington, DC	0.714	0.094	13.5 ± 10.7	56
University of Maryland	0.666	--	--	--
Read Street, Baltimore	0.639	0.095	9.1 ± 5.4	49
Essex, MD	0.606	0.074	10.7 ± 5.4	28
Linden, NJ	0.961	0.059	23.9 ± 28.5 (16.2 ± 9.9)* (n = 61)	65
Newark, NJ	0.510	--	--	--
East Boston, MA	0.777	0.063	31.3 ± 79.7 (9.6 ± 6.3)* (n = 61)	78
Watertown, MA	0.352	--	--	--

* Certain high ratios eliminated for these calculations (see text).

The high ratios in Linden, NJ and East Boston, MA stand out, including the high standard deviations for data from these two sites. Since there is a petroleum refinery near the Linden site and a tank farm upwind (on certain days) of the East Boston site, the individual sample data were inspected to determine if more information could be gleaned from an analysis of less source-specific data.

Data from the Linden site show that two of the individual sum-of-species/NO_X ratios were above 134, and two others were above 80. In three of these four, the NO_X concentrations seemed very low (around 0.01-0.02 ppm), while the fourth had a sum-of-species concentration of more than 6.8 ppm C. Eliminating these questionable samples from the calculations reduces the sum-of-species/NO_Xratio for the Linden site to 16.2 $\pm$ 9.9 (n = 61), which is a ratio comparable to the other cities. On the other hand, it should be borne in mind that relatively large sum-of-species /NO_X ratios may be characteristic of highly industrialized areas.

Similarly, eliminating from the calculations those samples collected at the East Boston site during which the wind was from the direction of the tank farm gives a sum-of-species/NO_X ratio of 9.6 $\pm$ 6.3 (n = 61). This ratio is now in line with the other three sites in Table 8.

All data pairs were used for the calculations of Table 8, except those noted in the above two paragraphs. A more informative number is the sum-of-species/NO_X ratio on those days which experienced high ozone values later in the day at a downwind monitoring site (although it is recognized that other factors - mainly meteorological - also contribute to the formation of ozone). Using only these days of high ozone for the HC/NO_X ratio calculation eliminates periods when other factors (low temperature, cloud cover, adverse winds, etc.)

mitigate against ozone formation. Table 9 summarizes these data for all days on which an ozone concentration of 0.15 ppm, or greater, was recorded at a site within about 50 miles of the station where hydrocarbon and NO_x concentrations were measured. Table 9 also shows NMOC/NO_x ratios on high ozone days together with sum-of-species/NO_x and NMOC/NO_x ratios (from Table 8) on all days of observation. These NMOC/NO_x ratios were included in order to gain some impression of the differences a photochemical modeler (particularly with the EKMA) might find if he had only continuous NMOC instrument data to use.

The comparison of ratios on high ozone days does not seem as bad as does the comparison of ratios on all days. But, as stated earlier, statistical summaries should be used with great caution. If the individual data which make up these summaries are examined, the photochemical modeler would have reason to be concerned. Table 10 shows the sum-of-species, NMOC, and NO_x data on high ozone days during the NECRMP in Essex, MD together with the calculated ratios. A similar picture of the data is seen at all the other sites on high ozone days: there is frequently a great difference between sum-of-species/NO_x ratios and NMOC/NO_x ratios on days of interest to photochemical modelers. Clearly, more work is needed to reconcile the differences in data obtained from these two measurement techniques.

More can be gleaned from Table 9. Both the Essex and the Linden sites show sum-of-species/NO_x ratios quite different from the other three sites on high ozone days. These two sites would probably be classified as industrial sites on the basis of these data, both exhibiting an excess of hydrocarbon as compared with sites influenced mostly by automobile exhaust. This is borne out to some extent by the data in Table 6. There, it is estimated that only 54 percent of the measured hydrocarbons come from automobiles at the Essex site,

Table 9

Hydrocarbon/NO_x Ratios On High Ozone Days

Site	Ratio - High O_3 Days						Ratio - All Days					
	Sum-of-Species/NO_x			NMOC/NO_x			Sum-of-Species/NO_x			NMOC/NO_x		
	mean	σ	n	mean	σ	n	mean	σ	n	mean	σ	n
West End Library, Washington, DC	7.8	3.6	16	6.6	2.8	16	13.5	10.7	56	6.2	3.2	51
Read Street, Baltimore, MD	6.9	2.7	15	10.8	5.3	15	9.1	5.4	49	12.9	12.6	36
Essex, MD	12.8	8.9	10	12.0	7.3	10	10.7	5.4	28	11.1	5.7	27
Linden, NJ	13.5	5.6	33	11.6	10.5	29	23.9	28.5	65	14.2	20.2	60
East Boston, MA	6.0	0.5	4	7.4	0.9	4	31.3	79.7	78	19.6	29.1	73

Table 10

Sum-of-Species/NO_X And NMOC/NO_X On High Ozone Days In Essex, MD

Date	Sampling Time	Sum-of-Species ppmC	NMOC ppmC	NO_X ppmC	Sum-of-Species/NO_X Ratio	NMOC/NO_X Ratio	Ozone ppm
8/14	6-7	0.676	0.90	0.099	6.8	9.1	0.15
	8-9	0.593	0.78	0.120	4.9	6.5	
8/24	6-7	0.339	0.27	0.050	6.8	5.4	0.15
	8-9	0.236	0.18	0.013	18.2	13.8	
8/26	6-7	0.944	1.14	0.081	11.6	14.1	0.17
	8-9	0.793	1.17	0.106	7.5	11.0	
8/27	6-7	2.755	2.22	0.159	17.3	14.0	0.18
	8-9	0.984	0.60	0.137	7.2	4.4	
8/29	6-7	0.516	0.45	0.039	13.2	11.5	0.15
	8-9	0.552	0.48	0.016	34.5	30.0	

and only 33 percent at the Linden site. In contrast, hydrocarbons at the West End Library site were estimated as 74 percent automotive; the University of Maryland site also as 74 percent; Read St, Baltimore as 72 percent; and Newark, NJ as 73 percent. The two Boston sites were lower, particularly the East Boston site, where the estimation was probably heavily influenced by the few days when wind was from the direction of a source to the northwest.

4. Aircraft and Roving Van Sample Data

Only a few words need be said about samples collected by aircraft and roving vans. The data user would be well advised to examine the individual sample data, rather than summaries of the data. Because of the frequently very low concentrations of species in the samples (80 percent of the aircraft samples and 20 percent of the roving van samples had less than 100 ppb C total hydrocarbon), there is great opportunity for large errors in the group fractions. The unidentified components of some aircraft samples exceed 50 percent of the total. Some aircraft samples are reported to have no olefins, others no aromatics. In some samples, the sum of aromatics is reported to exceed the sum of paraffins.

The roving van data, too, must be examined carefully. It may be that the chromatograph was not functioning properly during analysis of some of the samples, because some of the most commonly occurring aromatics are reported missing. In other samples, ethylene is the only reported olefin. These results are very unusual, and the user must assure himself, by using other information or by consulting chemists, that the reported numbers are reliable.

Summary

1. During the 1980 Northeast Corridor Regional Modeling Project (NECRMP) about 1120 ambient air samples were collected and analyzed from surface sites and by aircraft for hydrocarbon species analysis. The surface sites were in the metropolitan areas of Washington, DC, Baltimore, New York-New Jersey, and Boston.

2. An extensive QA/QC program carried out during the NECRMP showed that five participating laboratories could analyze standard gases repetitively with a coefficient of variation of 5-7 percent. C-2 to C-5 compounds in a standard mixture of 14 compounds were measured repetitively with a coefficient of variation of 15 percent. Heavy hydrocarbons, particularly aromatic species, were measured with much less precision. Round-robin samples (ambient samples collected and analyzed by one participating laboratory, then sent for consecutive analyses by all other participants) showed reasonable agreement ($\pm$ 15 percent coefficient of variation) as to total concentration of hydrocarbons in the samples, but there was frequent disagreement as to identification and quantity of some of the minor species.

3. There was an average of 16 percent (in concentration) of unidentified species in the 619 surface samples analyzed. Aircraft samples averaged much more. About 45 percent of the surface samples analyzed at six of the eight sites evidenced anomalous ethylene concentrations or ethylene contamination.

4. Data users are advised to examine data from the individual samples, rather than rely on summary statistics of those data, because of the characteristics summarized in "3" above.

5. Mobile source contribution to the observed hydrocarbon at each site was estimated. Automobiles were calculated to contribute more than 70 percent of the hydrocarbons at four of the eight sites, but the method of estimation is difficult to verify, thus the reliability of the estimate may have limited usefulness.

6. A comparison of sum-of-species concentration with data from collocated NMOC continuous analyzers was made. There was reasonable agreement among the data points at only one of the seven sites having NMOC instruments. Assuming sum-of-species data are more uniformly reliable than data from the continuous instruments, these latter did not agree very well with GC measurements under field conditions on this project. Even at sum-of-species concentrations exceeding 0.5 ppm C, there was reasonably good agreement between the two methods at only three of the sites.

7. Sum-of-species/NO_x ratios were computed for sites where NO_x instruments were collocated with species samplers. Computations were made, both for all days and for those days when ozone concentrations greater than 0.15 ppm were recorded downwind of the metropolitan areas. The average of this latter ratio differs by more than 2-fold between certain cities, but the number of high-ozone days was relatively few, and thus the statistical significance of the data is open to question.

8. Sum-of-species/NO_x ratios were compared with NMOC/NO_x ratios, both on high-ozone days and on all days for which data were available. The average of the ratios by these two methods on high ozone days seem to agree better than the average of the ratios on all days for which data are available, but comparison of individual day's ratios show large differences.

9. The fractions of species groups (olefins, paraffins, and aromatics) were shown to be fairly uniform among the four metropolitan areas.

References

1. W. A. Lonneman, "Ozone and Hydrocarbon Related Measurements in Recent Oxidant Transport Studies," International Conference on Photochemical Oxidant Pollution and Its Control, Proceedings: Volume I, EPA-600/3-77-001a, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, January 1977, pages 211-223.

2. H. H. Westberg, R. A. Rasmussen, and M. Holdren, "Gas Chromatographic Analysis of Ambient Air for Light Hydrocarbons Using a Chemically Bonded Stationary Phase," Anal. Chem. 46 (1974).

3. W. A. Lonneman, S. L. Kopczynski, P. E. Darley, and F. D. Sutterfield, "Hydrocarbon Composition of Urban Air Pollution," Env. Sci. Tech. 8(3) 229 (1974).

4. H. Westberg, K. Allwine, and E. Robinson, "Measurement of Light Hydrocarbons and Oxidant Transport, Houston Area, 1976," EPA-600/3-78-062, July 1978.

5. R. L. Seila, "Nonurban Hydrocarbon Concentrations in Ambient Air North of Houston, Texas," EPA-600/3-79-010, February 1979.

6. H. H. Westberg, W. A. Lonneman, and M. Holdren, "Analysis of Individual Hydrocarbon Species in Ambient Atmospheres - Techniques and Data Validity," American Chemical Society Annual Meeting, Kansas City, MO, September 1982.

7. W. A. Lonneman --- to be published.

8. F. Black and L. High, "Automotive Hydrocarbon Emission Patterns and the Measurement of Nonmethane Hydrocarbon Emission Rates," Paper 77-0144, International Automotive Engineering Congress and Exposition, Detroit, February 28 - March 4, 1977.

9. W. A. Lonneman, G. R. Namie, and J. J. Bufalini, "Hydrocarbons in Houston Air," EPA-600/3-79-018, February 1979.

10. F. W. Sexton, R. M. Michie, Jr., F. F. McElroy, and V. L. Thompson, "A Comparative Evaluation of Seven Automated Ambient Nonmethane Organic Compound Analyzers," EPA-600/54-82-046, August 1982, Environmental Monitoring Systems Laboratory, Research Research Triangle Park, NC 27711.

11. W. A. Lonneman, Lincoln Tunnel Studies, 1982 --- unpublished data.

12. L. R. Rechner, "Survey of Users of the EPA Reference Method for Measurement of Nonmethane Hydrocarbons in Ambient Air," EPA-650/4-75-008, December 1974; F. F. McElroy and V. L. Thompson, "Hydrocarbon Measurement Discrepancies Among Various Analyzers Using Flame-Ionization Detectors," EPA-600/4-75-010, September 1975; J. W. Harrison, M. L. Timmons, R. B. Denyszyn, and C. E. Decker, "Evaluation of the EPA Reference Method for the Measurement of Nonmethane Hydrocarbons - Final Report," EPA-600/4-77-033, June 1977.

13. J. R. Martinez, F. L. Ludwig, and C. Maxwell, "1978 Houston Oxidant Modeling Study, Volume I: Data Evaluation and Analysis," March 1982, (SRI Project No. 7938) p. 57ff.